RNA (1996), 2:201–212. Cambridge University Press. Printed in the USA.

A central pseudoknotted three-way junction imposes tRNA-like mimicry and the orientation of three 5′ upstream pseudoknots in the 3′ terminus of tobacco mosaic virus RNA

BRICE FELDEN, CATHERINE FLORENTZ, RICHARD GIEGÉ, and ERIC WESTHOF

Unité Propre de Recherche "Structure des Macromolécules Biologiques et Mécanismes de Reconnaissance," Institut de Biologie Moléculaire et Cellulaire du Centre National de la Recherche Scientifique, 15 rue René Descartes, 67084 Strasbourg-Cedex, France

ABSTRACT

A three-dimensional model of the histidylable 3′-terminal tRNA-like domain of tobacco mosaic virus RNA is proposed on the basis of a comparative structural analysis, chemical and enzymatic probing, combined with graphical modeling of three RNA constructs of increasing size (38, 108, and 182 nt) derived from the 3′-terminal viral RNA sequence. The comparison between the probing patterns of the three RNAs allowed the determination of the relative orientation of these structural domains in the full-length viral tRNA-like structure. Modeling data indicate that only one of the two possible isomers of the three-way junction located at a central position of the tRNA-like domain is in agreement with structural data. Interestingly, this isomer gives rise to a molecule bearing a structural mimicry with the L-shape of canonical tRNAs. A pseudoknotted acceptor branch containing a T-like loop is located perpendicularly to an anticodon-like branch. Moreover, a single-stranded RNA stretch belonging to the pseudoknotted central core mimics a D-like loop and it is proposed that it interacts via two conserved guanosines with nucleotides of the T-like loop as found in canonical tRNAs. This model is valid for the 3′ noncoding regions of tobamoviral RNAs as well as for the tRNA-like domain of the satellite tobacco mosaic virus RNA. All three molecules are substrates for yeast HisRS; however, whereas the complete viral genome is required for optimal histidylation capacities, both charging levels and affinity constants are decreased for the three RNA transcripts, suggesting that additional contacts located outside the tRNA-like domain are needed for an optimal aminoacylation process.

Keywords: minimalist tRNA-like structures; modeling; probing; three-dimensional structure; TMV; tRNA-like structures; viral RNA

INTRODUCTION

For several decades, the structure of the tobacco mosaic virus has been studied extensively (for reviews, see Holmes, 1984; Stubbs, 1984). However, despite several structural studies performed on the viral capsid, as well as on its interaction with the genomic RNA (e.g., Bloomer et al., 1978; Namba et al., 1989), only a few structural investigations were focused on the RNA itself, and these studies were restricted to its 5′ and 3′ ends. Based on the primary sequence of the viral genome (Guilley et al., 1979; Goelet et al., 1982), motifs responsible for enhancing translation were identified in the 5′ leader (Gallie & Walbot, 1992), whereas a pseudoknot-rich domain followed by a tRNA-like structure was found at the 3′ end (for review, see ten Dam et al., 1992). This particular 3′ domain, chargeable with a specific amino acid (Öberg & Philipson, 1972), increases both translational efficiency and RNA stability (Takamatsu et al., 1990; Gallie et al., 1991; Leathers et al., 1993). A model for the secondary structure of the TMV RNA 3′ end was first proposed by comparison with the TYMV RNA 3′ end (Rietveld et al., 1983). This was confirmed and further detailed by solution-probing data

Reprint requests to: Eric Westhof, Institut de Biologie Moléculaire et Cellulaire du Centre National de la Recherche Scientifique, 15 rue René Descartes, 67084 Strasbourg-Cedex, France; e-mail: westhof@bmc.u-strasbg.fr.

Abbreviations: TMV, tobacco mosaic virus; STMV, satellite tobacco mosaic virus; HDV, hepatitis delat virus; DMS, dimethylsulfate; DEPC, diethylpyrocarbonate; HisRS, histidyl-tRNA synthetase.

(Rietveld et al., 1984; Van Belkum et al., 1985; Garcia-Arenal, 1988), and artist drawings of the tertiary folding of the TMV tRNA-like domain for various viral strains were proposed. In the derived secondary structure, three major structural domains are connected by a central core (Fig. 1). A pseudoknotted domain D1 mimics a tRNA acceptor branch, an anticodon-like domain D2 is analogous to a tRNA anticodon branch, and a 5′ upstream domain D3 contains three pseudoknots. The pseudoknots of domains D1, D2, and D3 all belong to type (ii) (Westhof & Jaeger, 1992). These three structural domains are connected to a central core C, which comprises a three-way RNA junction in which two of the RNA helices connected by stretches of single-stranded RNA form a pseudoknot of type (i) (Westhof & Jaeger, 1992) with the L2 region comprising itself a helix. The conformation of this central pseudoknotted core is crucial for both the topology and the folding of the TMV tRNA-like domain because it imposes the orientations of the three adjacent domains D1, D2, and D3. Previous structural studies were unable to discriminate between several hypotheses concerning the folding of this central core (reviewed in Mans et al., 1991).

In this paper, we first confirm the secondary structure of the TMV tRNA-like domain (vulgare strain) by collecting additional enzymatic and chemical probing data, and then derive a refined three-dimensional model of the RNA structure by computer modeling. Moreover, this structural study was undertaken to define the tertiary folding of the central core C and to discriminate between the earlier proposals. To address these issues, three RNA constructs of increasing size were designed: an RNA transcript of 38 nt analogous to domain D1, which mimics a tRNA minihelix; a 108-nt long RNA corresponding to the TMV tRNA-like core (domains D1 and D2); and, finally, an RNA transcript of 182 nt corresponding to the entire TMV tRNA-like structure (domains D1, D2, and D3). A comparative analysis of their reactivities toward chemical and enzymatic probes allowed definition of the correct orientation of these three domains relative to the central core.

The modeling results show that two alternate conformers are plausible for the folding of the three-way RNA junction, but both probing and functional data are only compatible with one of the two putative conformational isomers. Interestingly, only one isomer al-

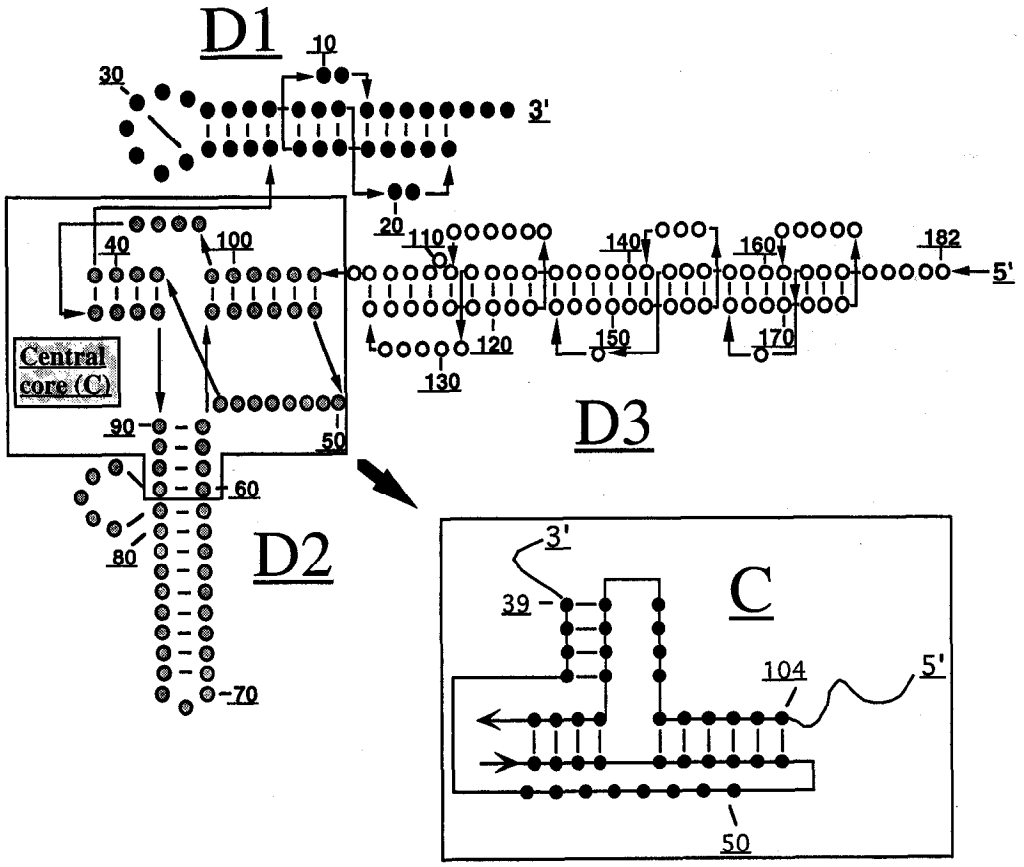

FIGURE 1. Schematic representation of the TMV tRNA-like secondary structure on the basis of previous structural studies emphasizing the three structural domains D1 (black dots), D2 (grey dots), and D3 (white dots), all connected together by a central core C (boxed). The boxed region, or central core C, is redrawn differently to illustrate the pseudoknot topology. As usual in the tRNA-like field, the numbering starts at the 3′ end. Arrows follow the RNA strand from the 5′ to the 3′ end. D1 (38 nt), D1-D2 (108 nt), and D1-D2-D3 (182 nt) correspond to the three RNA constructs used for structural investigations. In the virus, the D1-D2-D3 domains are covalently linked to ~6,200 nt.

lows a perpendicular orientation of domains D1 and D2, thereby mimicking the two branches of canonical tRNAs. Moreover, the respective orientation of these two domains seems constrained by two long-range interactions similar to those existing between the D- and T-loops in canonical tRNAs. Because the TMV RNA is aminoacylable with histidine (Öberg & Philipson, 1972), the histidylation capacities of the three RNA constructs were measured using yeast histidyl-tRNA synthetase (HisRS) and compared to that of the entire TMV RNA.

RESULTS

Secondary structure of the 182 nucleotide long TMV tRNA-like domain

Whereas previous structural studies were performed on RNA fragments of more than 100 nt derived from the natural TMV tRNA-like domain (Rietveld et al., 1984; Van Belkum et al., 1985), the present study investigates the structure of three in vitro transcribed overlapping RNAs of increasing size derived from the 3′ end of the viral genome (38, 108, and 182 nt). Chemical (DMS and DEPC) as well as enzymatic (nuclease S1 and RNases V1 and T1) probes were applied. Typical modification patterns, obtained after chemical probing of the three labeled RNA transcripts, are displayed in Figure 2. The summary of all enzymatic and chemical probing data obtained on the three RNAs is given in Figure 3. Structural information is available for each nucleotide, as indicated on the secondary structure of the TMV tRNA-like domain proposed by Rietveld et al. (1984), and which includes several additional structural features derived from our structural data, namely an A30-U34 reverse-Hoogsteen base pair, the stacking of nt A18 on C17, and the location of the single-stranded RNA stretch A95–U98 in close proximity of the U28–U34 loop.

Despite a couple of differences, the present structural data are in good agreement with those reported previously for a 104-nt long tRNA-like fragment (Rietveld et al., 1984) and for the 204-nt long 3′ noncoding region of TMV RNA (vulgare strain) (Van Belkum et al., 1985). Indeed, V1 cuts are always observed in double-stranded regions of the proposed secondary structure and S1 cleavages were only found in loops or in single-stranded RNA stretches (see Fig. 3 for details). Moreover, our additional probing data concerning Watson–Crick positions of cytosines (atoms N-3) support most of the base pairings of the secondary structure. Indeed, as shown in Figure 3, cytosines involved in Watson–Crick base pairings within the secondary structure of the TMV tRNA-like domain are reactive only in denaturing conditions (one exception is C100 in the 182-nt long RNA transcript). In addition, Hoogsteen positions of purines located in helical domains are only reactive toward chemical probes when the RNA structure is denatured, another observation in favor of the proposed secondary structure. Moreover, the five pseudoknots of the tRNA-like structure proposed previously, one in domain D1, one in the central core C, and three in domain D3 (Rietveld et al., 1984), are validated by our additional probing data, and especially at both the Watson–Crick and Hoogsteen positions of the nucleotides involved in these tertiary interactions (see Fig. 3 for details). According to its reactivity toward chemicals, the 7-nt loop U28–U34 mimics a T-loop structure as found in canonical tRNAs with a A30–U34 reverse Hoogsteen base pair stacked on the last base pair of the adjacent helix, and nt U28 and A29 located outside from the loop.

Structural relationships between domains D1, D2 and D3

The mapping data of common structural domains within the three RNA structures of increasing size (RNA D1, RNA D1-D2, and RNA D1-D2-D3) have been compared in order to establish potential structural relationships between the different domains of the full-length molecule. Several differences can be observed between the shortest RNA fragment and the intermediate size. Two sets of V1 cuts, between C4 and G6 and between G11 and C13, are observed in RNA D1, but not in RNA D1-D2. On the other hand, RNA D1-D2 is cleaved weakly by RNase V1 between C16–C17, U20–G21, and G37–G38 residues, and by nuclease S1 between G31–U33 residues, but RNA D1 is not. Finally, in RNA D1-D2, four V1 cuts (between G21 and C25) are shifted by one residue to the 5′ side in RNA D1 (V1 cuts between C22 and C26). In contrast to RNA D1, domain D1 is in close proximity to the central core in RNA D1-D2. The proximity of domain D1 in RNA D1-D2 may account for the differences observed between the two molecules when comparing their probing patterns. Further, the disappearance of cuts at the 3′ end of D1 and the appearance of new cuts in D1-D2 on the complementary side indicate a destabilization of the pseudoknot in the longer molecules, possibly due to the constraints imposed by the short junction between D1 and D2. Interestingly, the probing pattern of domain D1 is identical between the intermediate and the longest RNA fragment (see Fig. 3 for details). These results suggest that no tertiary interaction is occurring between domains D1 and D3 which, therefore, appear to behave independently.

The probing patterns of domain D2 within the two RNA constructs D1-D2 and D1-D2-D3 have also been compared. Although a similar pattern is observed for nt U61–G86, corresponding to the part of the tRNA-like structure mimicking an anticodon stem-loop, a strong difference is noticed within the central core C (nt A39–C60 and G87–G104). Indeed, strong V1 cuts at residues G102, U103, and G104 occur only in the full-length

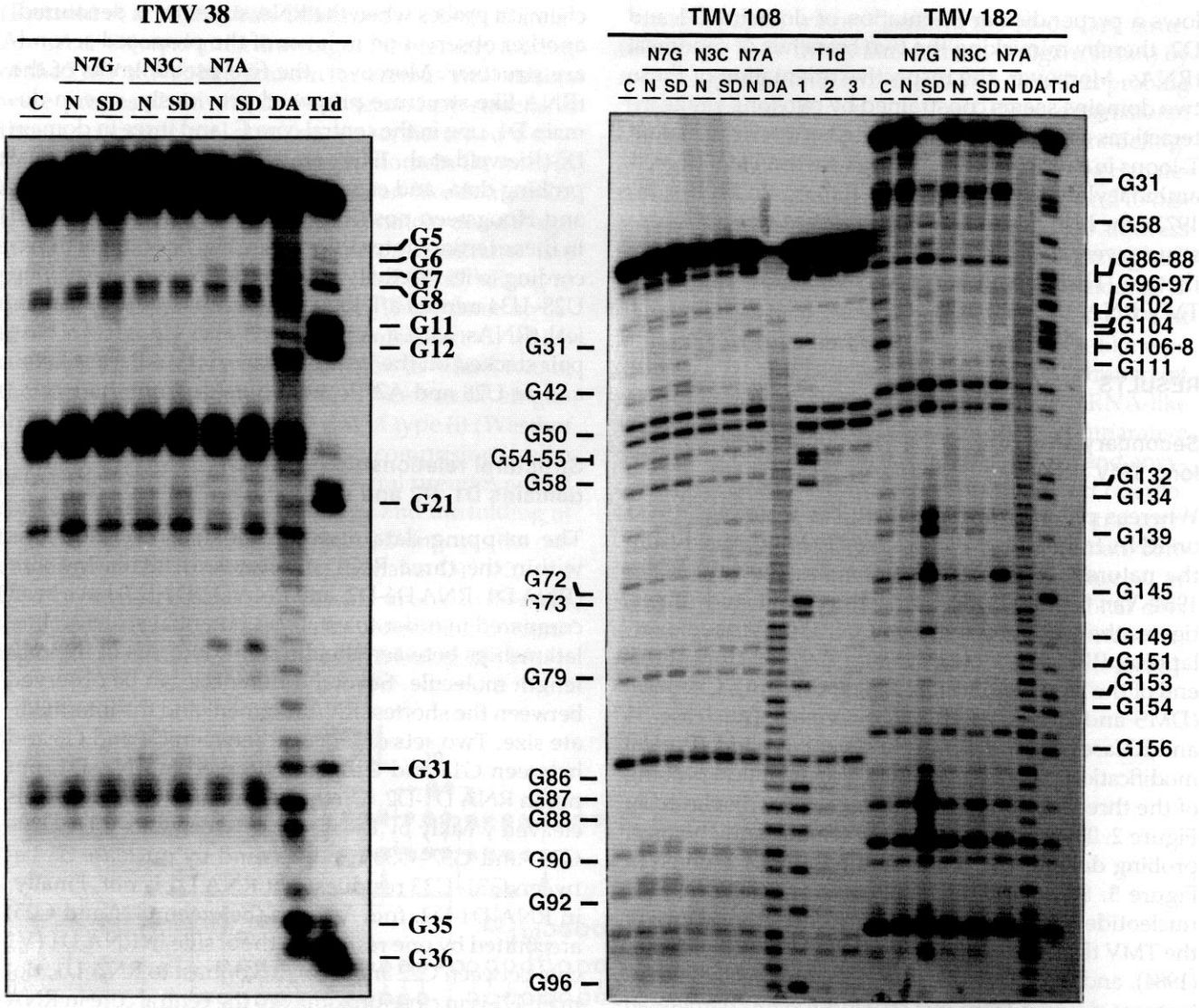

FIGURE 2. Mapping of N-7 positions in the purines and of N-3 positions in cytidines of the three RNA constructs (TMV 38, 108, and 182) by DMS (atoms N-3 of C and atoms N-7 of G) and DEPC (atoms N-7 of A). Autoradiograms of 20% (left) and 12% (right) polyacrylamide gels of cleavage products of 5′-labeled RNAs. Lanes C, incubation controls; lanes N, probing under native condition (4 min incubation with 1 μL of DMS diluted twice in ethanol, and 15 min with 10 μL of pure DEPC); lanes SD, probing under semi-denaturing conditions (4 min incubation with 1 μL of DMS diluted twice in ethanol, and 15 min with 5 μL of pure DEPC); lanes DA, formamide ladder; lanes T1d, guanine ladder with different enzyme concentrations (lanes 1–3). Sequencing tracks are numbered at each G residue.

RNA construct, indicating that domain D3 is needed for the occurrence of V1 cleavages in this region and, thus, for the stabilization of this helical domain.

Taken together, these comparative probing data between the three RNAs bring information for the construction of the model of the largest fragment and especially for the assembly of the secondary elements D1, D2, and D3 into a three-dimensional fold.

Modeling of the central core

The three-dimensional conformation of the central core C is crucial because it sets constraints on the three adjacent domains D1, D2, and D3. In previous structural studies, several alternate foldings were already proposed for this three-way RNA junction (reviewed in Mans et al., 1991). In the absence of biochemical data and with due consideration of the geometry and stereochemistry of RNAs, the pseudoknotted central core of the TMV tRNA-like domain can be constructed in two different ways, as shown in Figure 4. The two isomers possess the same topology, but opposite chiralities. The geometry of the three-way RNA junction is dictated by the chirality of the two connecting single strands, which impose the respective orientation of the three RNA helices. When those single-stranded stretches follow a

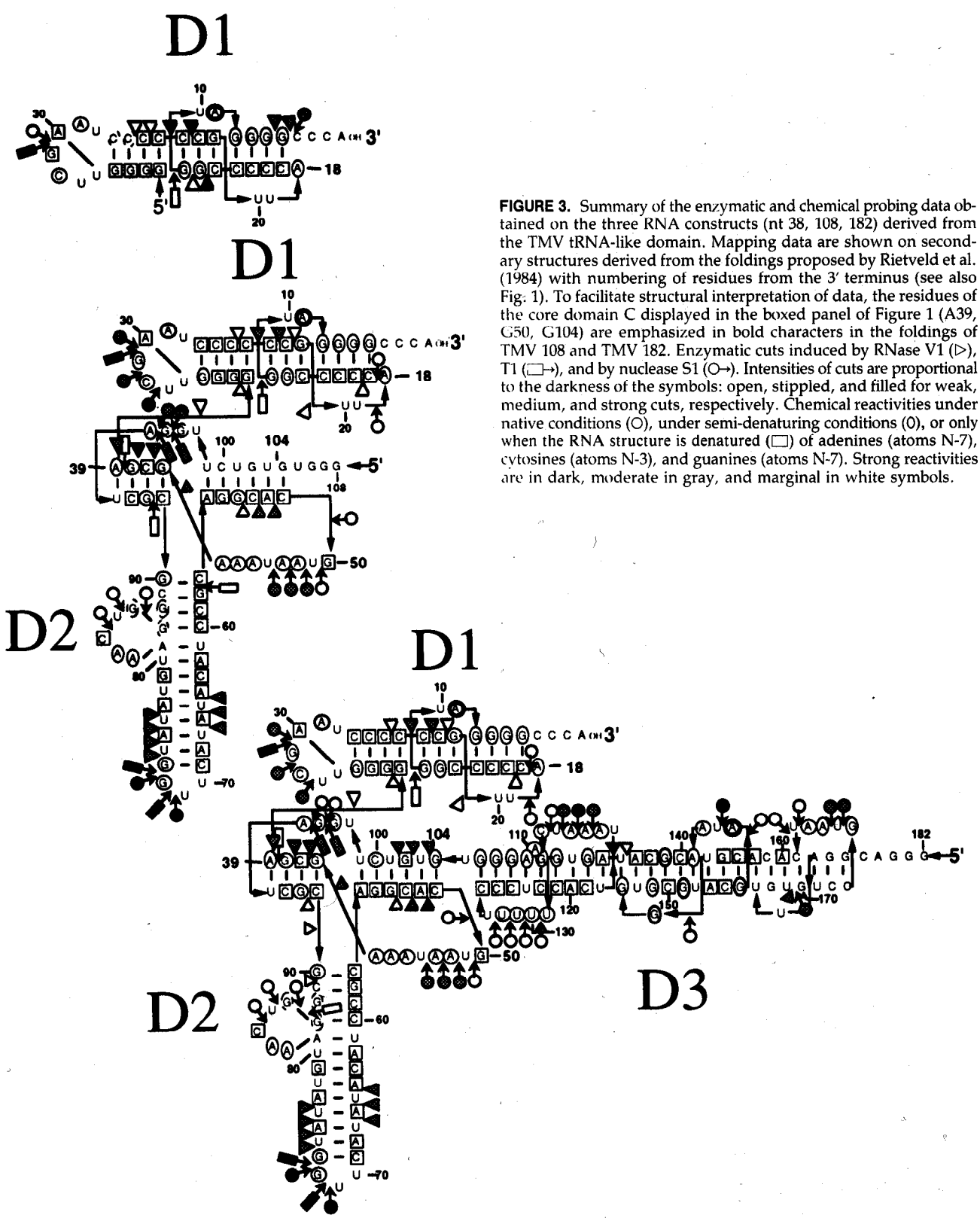

FIGURE 3. Summary of the enzymatic and chemical probing data obtained on the three RNA constructs (nt 38, 108, 182) derived from the TMV tRNA-like domain. Mapping data are shown on secondary structures derived from the foldings proposed by Rietveld et al. (1984) with numbering of residues from the 3′ terminus (see also Fig. 1). To facilitate structural interpretation of data, the residues of the core domain C displayed in the boxed panel of Figure 1 (A39, G50, G104) are emphasized in bold characters in the foldings of TMV 108 and TMV 182. Enzymatic cuts induced by RNase V1 (▷), T1 (□→), and by nuclease S1 (○→). Intensities of cuts are proportional to the darkness of the symbols: open, stippled, and filled for weak, medium, and strong cuts, respectively. Chemical reactivities under native conditions (○), under semi-denaturing conditions (◐), or only when the RNA structure is denatured (□) of adenines (atoms N-7), cytosines (atoms N-3), and guanines (atoms N-7). Strong reactivities are in dark, moderate in gray, and marginal in white symbols.

right-handed path, the "right-handed" isomer is obtained (Fig. 4, left panel). On the other hand, the "left-handed" isomer is obtained when the single-stranded stretches follow a left-handed path (Fig. 4, right panel).

Our probing data were sufficiently accurate to discriminate which one of the two conformations occurs predominantly in solution. In Figure 4, the white spheres (the two bottom panels) correspond to the

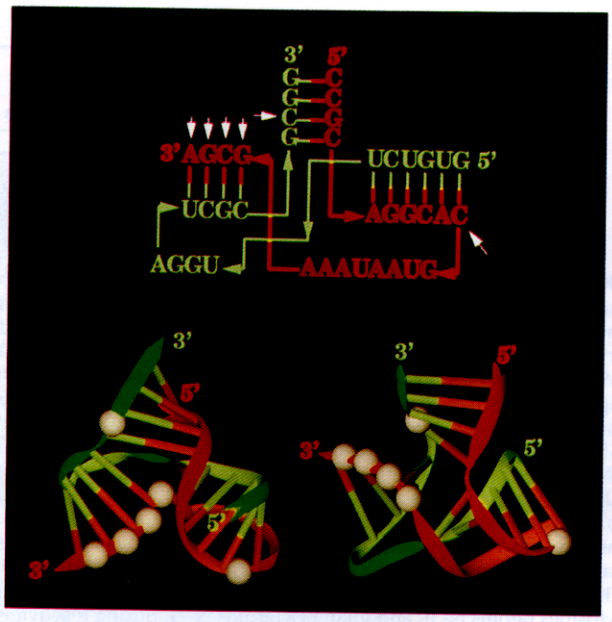

FIGURE 4. Secondary structure and graphical modeling constructions of two possible conformational isomers for the central core of the TMV tRNA-like domain. White arrows on the secondary structure (top) correspond to white spheres on the two models (bottom) and indicate the location of RNase V1 cleavages derived from the probing data. The structure on the left corresponds to the left-handed isomer and the one on the right to the right-handed isomer. According to the high sterical hindrance of the V1 enzyme, only the right-handed model of the central core (bottom right) accounts well for the cleavage pattern. For simplicity, one RNA strand is colored red and the other green on both secondary and tertiary structures of the three-way junction. Both 3′ and 5′ ends of each RNA strand are indicated.

white arrows (top panel) and indicate the exact location of RNase V1 cuts. Considering the high sterical hindrance of RNase V1, an enzyme of 15,900 Da, only the right-handed isomer is in agreement with the cleavage pattern. Indeed, two (those located at the 5′ side) of the four enzymatic cuts occurring in a stretch are not possible in the left-handed conformation because they are located in an inaccessible area due to the very close proximity of the adjacent RNA helices. These two V1 cuts are well explained in the right-handed conformation (see Fig. 4 for details). The other V1 cleavage sites were not able to discriminate between these two putative RNA foldings. Thus, the right-handed isomer is favored strongly in solution. According to this conformation of the central core, the respective orientation of the three structural domains D1, D2, and D3 was defined and a three-dimensional model rationalizing both structural and functional data was derived by computer modeling.

Structural comparisons: Is the RNA world right-handed?

In the two isomers of opposite chirality, there is no helix–helix stacking and the three-way junction adopts a Y-shaped conformation as already proposed for three-way DNA junctions (Duckett et al., 1990; Leontis et al., 1993). This is a major discrepancy with previous proposals for alternate foldings of this central core. Indeed, these models were only based on different stacking choices of the three RNA helices (reviewed in Mans et al., 1991).

A well-documented example of other three-way RNA junctions are the *Xenopus laevis* (Westhof et al., 1989) and *Escherichia coli* (Brunel et al., 1991) 5S ribosomal RNAs. Both 5S rRNAs fold into a Y-shaped structure in which two helices are quasi-colinear but, in that case, two of the three structural domains are connected by unpaired nucleotides involved in a precise array of hydrogen bonds based on triple interactions on the shallow groove side. The hammerhead ribozyme belongs to the three-way junctions also with two stacked helices, but with two connecting single-stranded stretches involved in nonclassical base pairing on the shallow groove side. Fluorescence experiments (Tuschl et al., 1994), coupled to modeling, led to an architecture similar to that obtained by X-ray crystallography (Pley et al., 1994; Scott et al., 1995), except for the relative orientation of the unstacked helix with respect to the stacked stem. In the central core of TMV, the three-way junction results from pseudoknotting between a bulge and 3′ dangling strand. Therefore, one does not expect similarities between the hammerhead ribozyme and the central core of TMV RNA. However, a striking and unexpected topological similarity exists (Fig. 5) between the C domain of TMV RNA and the catalytic region of

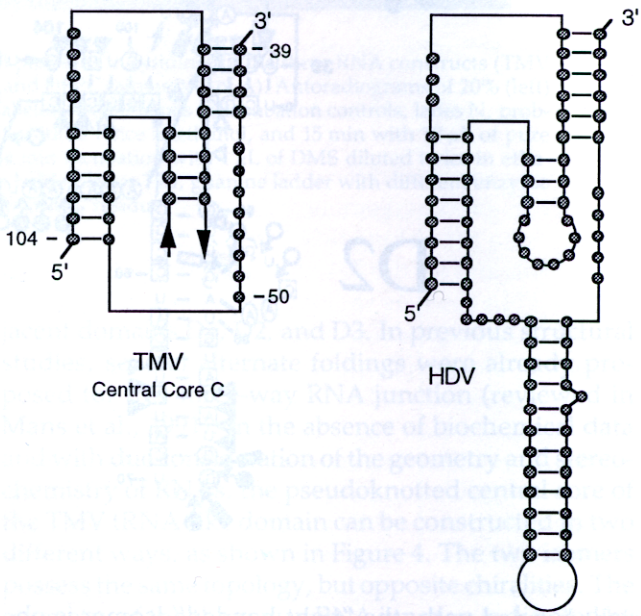

FIGURE 5. Two-dimensional drawings illustrating the topological analogy between the central domain C in TMV (left) and the catalytic domain of HDV (right) folded as a pseudoknot according to Perrotta and Been (1991).

hepatitis delta virus (HDV), which folds also into a pseudoknot (Perrotta & Been, 1991). In a proposed model for the HDV core (Tanner et al., 1994), the internal helix is short and subtends a hairpin loop carrying catalytically active residues, whereas in TMV, the internal helix leads to the variable and long domain D2. It is thus not possible to build both domains with a similar arrangement of helices.

For modeling purposes, it is worthwhile to compare three-way junctions with four-way junctions. In four-way junctions (Lilley & Clegg, 1993), there are four isomers: the crossed and noncrossed junctions lead to right-handed or left-handed arrangement of the stacked two arms forming the X-shaped conformation. In that instance, the right-handedness at the junction controls again the choice between crossed and noncrossed isomers. In the crossed junctions, the path of the joining backbone is left-handed and, in the noncrossed junction, it is right-handed (Krol et al., 1990). The most commonly observed conformer in RNA and DNA is the noncrossed right-handed arrangement of the four arms (Krol et al., 1990; Lilley & Clegg, 1993; Duckett et al., 1995).

The folding of the TMV tRNA-like domain

The three-dimensional folding of the 182-nt long RNA fragment, encompassing domains D1, D2, and D3, and corresponding to the TMV tRNA-like domain, is shown in Figure 6. Both N-7 reactivities of purines located between connecting helices and the comparative structural analysis of similar domains within the three RNA fragments were essential to connect the secondary structural elements into a three-dimensional fold. The topology of the RNA structure is as follows: D1 is perpendicular to D2 and D3 is directed in nearly an opposite direction to D1. This gives rise to a bent D1-D2 domain. When using the left-handed conformation of the central core, the topology of the tRNA-like domain is totally changed, giving rise to a global Y-shaped fold where the three domains D1, D2, and D3 are oriented at about 120° from one another (result not shown).

The conformations and proximity of domains D1 and C suggest that they are connected by two tertiary interactions C32–G96 and U33–G97, mimicking the G18–Ψ55 and G19–C56 found in canonical tRNAs between D- and T-loops, respectively (Rich & RajBhandary,

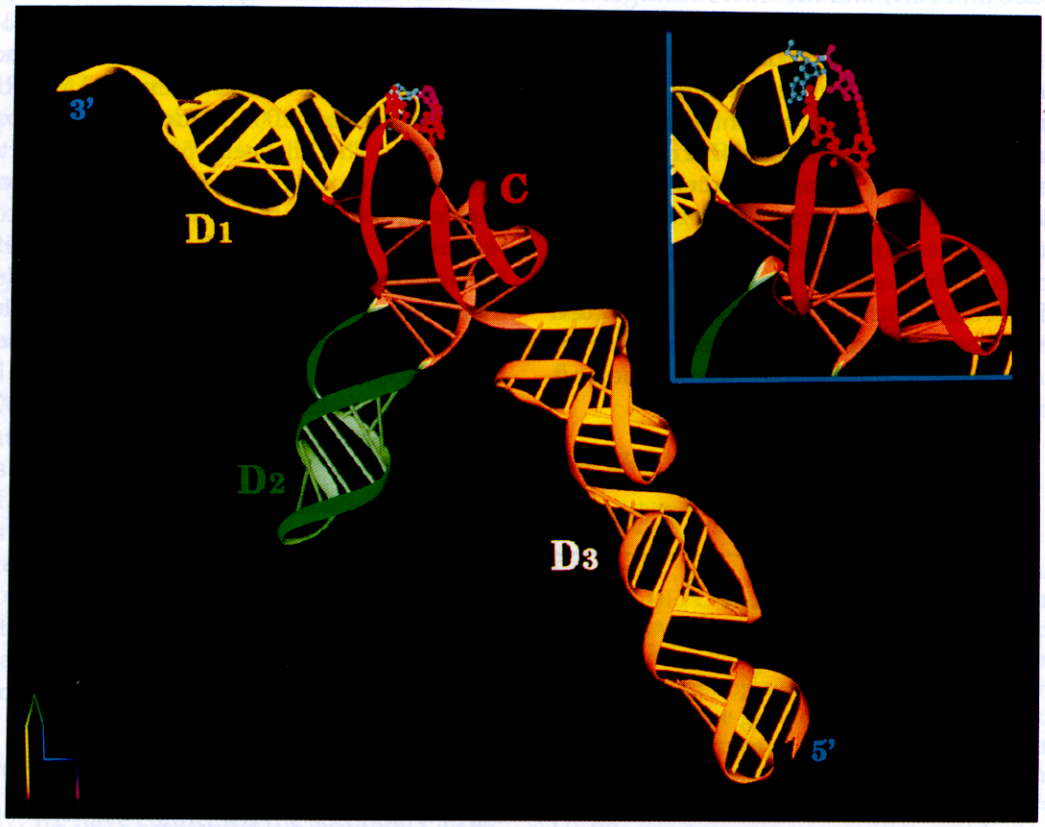

FIGURE 6. Color view of the three-dimensional model of the TMV tRNA-like domain corresponding to the RNA fragment of 182 nt and emphasizing the three structural domains D1, D2, and D3. Domains D1 (in yellow) and D2 (in green) are located perpendicularly and mimic the L-shape of canonical tRNAs. Domain D3 (in orange) is composed of three pseudoknots coaxially stacked. Note that the conformation of the central core C (in red) imposes the relative orientation of the three structural domains. In the insert, the base tertiary interactions between loops in the central core C and domain D1, which mimic the interactions between the D- and T-loops in tRNAs, are explicitly shown.

1976). The single-stranded RNA stretch (A95, G96, G97, U98) located in the pseudoknot of the central core indeed mimics a D-loop conformation. The chemical and enzymatic data do not support convincingly the proposed tertiary contacts. However, those tertiary pairs are exposed at the edge of the molecule so that access to chemicals and enzymes is not prevented. The equivalent positions in tRNAs are often found reactive (Romby et al., 1987). The occurrence of these two novel tertiary interactions would be strengthened by a strict conservation of these two guanosine residues G96 and G97 for other sequenced tobamoviral RNAs, namely TMV vulgare strain (Goelet et al., 1982), cucumber green mottle mosaic virus (Meshi et al., 1983), TMV tomato strain (Ohno et al., 1984), TMV U2-strain (Garcia-Arenal, 1988), pepper mild mottle virus strain S RNA (Avila-Rincon et al., 1989), tobacco mild green mosaic virus (Solis & Garcia-Arenal, 1990), and even in the case of the satellite tobacco mosaic virus RNA (Felden et al., 1994b), whereas flanking sequences are random. Based on the strong conservation of G96 and G97, Mans et al. (1991) already hypothesized about a relationship between the T-like loop and the A95–U98 regions. In the vulgare strain, domain D1 is connected directly to the central core by the phosphodiester bond between nt G38 and A39, and these two residues belong to two different helical domains. Moreover, the rotational freedom of this phosphodiester bond would be further restricted by the suggested tertiary interactions between domain D1 and the central core. Consequently, in the present three-dimensional model, the central core sets the orientation of D1 and enforces the spatial location of D3 by stacking interactions between the three pseudoknots and the adjacent helix of the three-way junction. However, the orientation of the anticodon-like domain with respect to the central core may differ from one viral strain to another because an internal RNA bulge varying largely in sequence and length is located in between.

Functional implication: Histidylation properties of the three RNAs derived from the TMV tRNA-like domain

The genomic RNA of TMV has been known to be histidylable at its 3′ end by HisRS from several origins for more than 20 years (Öberg & Philipson, 1972). The elements, within canonical histidine-accepting tRNAs, responsible for their specific aminoacylation have been defined in several instances (Himeno et al., 1989; Francklyn & Schimmel, 1990; Rudinger et al., 1994). The major identity element concerns base −1, located at the 5′ end of the tRNA. This is a feature only encountered in $tRNA^{His}$ (Steinberg et al., 1993), which finds its structural equivalent in the acceptor stem of the tRNA-like domain from turnip yellow mosaic virus RNA (Rudinger et al., 1992). An equivalent nucleotide is present in the pseudoknot of domain D1 of the TMV tRNA-like domain, namely nt 18, and is an excellent candidate responsible for the histidylation properties of the viral RNA. If this holds true, the shortest RNA fragment investigated here and mimicking a tRNA acceptor branch, should be an efficient substrate for yeast HisRS.

The explicit demonstration of the histidylation properties of this RNA is given in Table 1. Yeast HisRS is able to aminoacylate up to 40% of RNA D1. The presence of additional domains D2 and D3 increases the charging level to 66%. The entire TMV RNA can be charged to nearly 100% under the same experimental conditions. The Michaelis–Menten parameters K_m and V_{max} give a more quantitative estimation of the histidylation properties of the three molecules compared to full-length TMV RNA. The histidylation efficiency of D1, which corresponds to a minimalist RNA substrate mimicking a tRNA accepting branch, is high compared to that of the full-length TMV RNA genome because it is only about 100-fold less active ($L = 107$, where L is expressed by the ratio $(V_{max}/K_m)_{\text{TMV RNA}}/(V_{max}/K_m)_{\text{D1}}$).

TABLE 1. Kinetic parameters of histidylation of the three RNAs (nts 38, 108, 182) derived from the TMV tRNA-like domain.[a]

RNA	Aminoacylation plateaus[b] (%)	K_m (nM)	V_{max} (arbitrary units)	V_{max}/K_m (arbitrary units)	L (relative to TMV RNA)
D1	40	4,000	140	0.035	107
D1-D2	56	3,900	400	0.10	38
D1-D2-D3	66	3,100	650	0.21	18
TMV RNA	98	90	340	3.76	1

[a] Aminoacylation levels were obtained with equivalent amounts of enzyme and correspond to optimal plateaus. L value corresponds to loss of specificity and is expressed by the ratio $(V_{max}/K_m)_{\text{TMV RNA}}/(V_{max}/K_m)_{\text{D1 or D1-D2 or D1-D2-D3}}$.

[b] It is recalled that aminoacylation plateaus of less than 100% do not necessarily reflect existence of inactive molecules. Decreased aminoacylation levels, accompanied by decreased charging efficiencies, reflect the existence of the equilibrium between the aminoacylation reaction and those reactions leading to the deacylation of the charged RNAs [for details on the incomplete charging of tRNA-like molecules see Giegé et al. (1978) and references therein].

For comparison, the aminoacylation of the minihelix derived of yeast $tRNA^{Asp}$ by yeast AspRS is 9,000-fold less efficient than the aminoacylation of the corresponding entire tRNA (Frugier et al., 1994). Interestingly, if the anticodon-like branch D2 is added to RNA D1, the histidylation efficiency is increased threefold ($L = 38$). Consequently, domain D2 may interact with yeast HisRS during the aminoacylation process, as already suggested by nuclease protection experiments of the TMV RNA by HisRS (Garcia-Arenal, 1988). Moreover, when domain D3 is added to RNA D1-D2, the specificity of the aminoacylation reaction is again increased by a factor of two ($L = 18$). Additional contacts between the viral RNA and the enzyme may occur outside the strict tRNA-like core, which comprises solely domains D1 and D2. Finally, whereas the V_{max} values are comparable between the four RNAs, the affinity of the entire TMV RNA genome for yeast HisRS is much higher than that for the three RNA fragments (90 nM versus 3,000–4,000 nM). These results are not understood clearly, but it may be proposed that additional contacts between the viral genome and yeast HisRS located outside the tRNA-like domain (RNA D1-D2-D3) increase the affinity between the two interacting molecules.

Taken together, these functional data indicate that the three domains D1, D2, and D3 are of importance for the aminoacylation properties of the viral tRNA-like domain. Additional functional data obtained on viral RNA mutants will be essential to define the full set of recognition elements required for the optimal aminoacylation capacities of the TMV RNA.

In a previous study, Joshi et al. (1985) have searched for the minimal size of the RNA needed for histidylation activity. This was performed using electrophoretic methods to isolate 3′ fragments labeled by aminoacylation by a partially purified sheep liver HisRS of mildly hydrolyzed viral RNA. They came to the conclusion that all RNA molecules longer than about 95 nt can be aminoacylated. These results are at variance with the conclusions of this work, showing that a much shorter fragment of 38 nt can be histidylated. The discrepancy, however, is only apparent because the previous experiments were done with synthetase preparations from different sources and of different purities. It is quite likely that the shorter fragments, charged with a decreased catalytic efficiency, could not be revealed under the former experimental conditions.

CONCLUSIONS

In this study, we have confirmed the secondary structure of the TMV tRNA-like domain by collecting additional solution-probing data and we derived a refined three-dimensional model of the RNA structure by computer modeling. We showed that the conformation of a three-way RNA junction located at a central position within the structure imposes the folding of the entire tRNA-like domain. Two conformers of opposite chiralities are geometrically plausible, a right-handed and a left-handed isomer, but only the right-handed conformation is in agreement with the probing data. Moreover, our functional data indicate that the full-length TMV RNA is histidylable up to 100%. Because only the selected isomer allows mimicking of the folding of a canonical tRNA, both functional and structural results are in line with the absence of any equilibrium between the two isomers.

A refined three-dimensional model, in which the orientation of the three structural domains D1, D2, and D3 is constrained by the conformation of the central core, is proposed. This model is valid for the 3′ noncoding domains of tobamoviral RNAs as well as for the tRNA-like domain of STMV RNA. Domains D1 and D2 are located perpendicularly and mimic the L-shaped fold of canonical tRNAs. Interestingly, our data suggest that, in the case of the viral tRNA-like domain, a D-loop is mimicked by a different structural feature than that found in transfer RNAs, namely a single-stranded RNA stretch (A95–U98) belonging to the pseudoknotted central core. Two tertiary interactions, occurring between the D-like and T-like loops, and involving four conserved residues among the tobamoviral RNAs, are proposed. Consequently, as has already been established for the TYMV (Dumas et al., 1987) and for the BMV tRNA-like domains (Felden et al., 1994a), similar functions, namely the aminoacylation properties of tRNAs and of plant viral RNAs, are linked to similar tertiary foldings, but arise from different secondary structural elements. In contrast, although a topological analogy exists between the pseudoknotted three-way junctions of the central domain C in TMV and of the catalytic domain of HDV, the three-dimensional models proposed for TMV and HDV differ and there is no reported functional activity.

MATERIALS AND METHODS

Chemicals and enzymes

DMS was from Aldrich-Chimie (St Quentin-Fallavier, France), DEPC and hydrazine from Sigma (St. Louis, Missouri, USA), and aniline from Merck (Darmstadt, Germany). Nucleotides were from Boehringer-Mannheim (Meylan, France). Rotiphorèse Gel 40 solution of acrylamide and *N,N′*-methylene-bis-acrylamide was from Carl Roth GmbH (Karlsruhe, Germany). Radioactive [^{32}P]pCp at 3,000 Ci/mol and [γ-^{32}P]ATP at 3,200 Ci/mol were from Amersham (Les Ulis, France), as was L-[^{3}H] histidine (58 Ci/mmol). Yeast HisRS was an enzyme preparation enriched by chromatographies on DEAE-cellulose, hydroxyapatite, and phosphocellulose (Mengual, 1977). Restriction enzyme *Eco*T22I was purchased from United States Biochemical Corporation (Cleveland, Ohio, USA). The nucleases S1, T1, and V1 were from Pharmacia (Paris, France). Phage T4 polynucleotide ki-

nase was from Amersham (Les Ulis, France), snake venom phosphodiesterase was from Worthington (Freehold, New Jersey, USA), bacterial alkaline phosphatase was from Appligène (Strasbourg, France). Bacteriophage T7 RNA polymerase was prepared according to Wyatt et al. (1991). Total yeast tRNA, used as a carrier RNA to supplement labeled viral RNA fragments, was from Boehringer-Mannheim (Meylan, France).

Preparation of RNAs and aminoacylation assays

The RNA constructs are derived from the wild-type sequence of TMV RNA (vulgare strain). Their cloning downstream from the T7 RNA polymerase promoter was done as described earlier (Perret et al., 1990) using synthetic oligonucleotides. Plasmids were linearized by *Eco*T22I restriction nuclease before transcription so that transcripts will end with the 3'-terminal CCA triplet. In vitro transcription was performed as described (Felden et al., 1994a). The three RNA transcripts were separated from nonincorporated nucleotides and DNA fragments by electrophoresis on denaturing gels. Appropriate bands were electroeluted and pure transcripts were recovered by ethanol precipitation. Because all RNA variants were handled strictly under identical conditions, we believe that they are in the same conformation in their functional core (see also the legend to Table 1). TMV virus (Strasbourg strain) was a kind gift of J. Witz (Strasbourg) and was prepared as described (Von Wechmar & Van Regenmortel, 1970). Viral RNA was extracted from its capsid by two phenol extractions and recovered by ethanol precipitation. Spectrophotometric measurements were made to determine the concentration of the RNA, assuming that 1 absorbance unit at 260 nm corresponds to 40 $\mu g \cdot mL^{-1}$ RNA in a 1-cm-pathlength cell.

Aminoacylation reactions were performed on the three RNA transcripts and on the full-length TMV RNA. Charging plateaus and kinetic parameters were done at 30 °C in a medium containing 25 mM Tris-HCl, pH 7.8, 15 mM $MgCl_2$, 7.5 mM ATP, 50 μM L-[^{3}H] histidine (58 Ci/mmol), and the required concentrations of RNA and yeast HisRS. Aliquots were spotted on 3MM Whatman papers at four different times and precipitated in 10% (v/v) trichloroacetic acid. Kinetic parameters (K_m and V_{max}) were determined from Lineweaver-Burk plots. Before aminoacylation measurements, RNA molecules were renatured by a heating/cooling treatment (Frugier et al., 1994).

Structural mapping procedures

Labeling of the 5' ends of the three RNAs was performed with [γ-^{32}P]ATP and phage T4 polynucleotide kinase on RNA dephosphorylated previously with alkaline phosphatase (Silberklang et al., 1977). Labeling at the 3' end was done by ligation of [γ-^{32}P]pCp with T4 RNA ligase (England & Uhlenbeck, 1978). Before either enzymatic digestions or chemical modifications, the labeled RNAs were heated up to 50 °C for 5 min and slowly cooled down at room temperature for 20 min.

Digestions with the various nucleases (V1, S1, T1,) were done at 20 °C as described (Felden et al., 1994a) on labeled transcripts (30,000–60,000 cpm depending of the experiments) supplemented with 1 μg of total tRNA. The following amounts of nucleases were added: 1.25 10^{-2} unit of RNase T1, 0.1 unit of RNase V1, and 83 units of nuclease S1. Incubation times were 10 min for V1 and T1 cleavages and 5 min for S1 mapping.

Modification of N-3 atoms of cytosine and N-7 atoms of guanine residues by DMS and of N-7 positions in adenine by DEPC were done at 20 °C according to Peattie and Gilbert (1980). Concentration of chemicals and incubation times have been optimized for the three RNAs. Reaction mixtures of 200 μL contain the appropriate buffer, the labeled transcripts (60,000 cpm) supplemented with 10 μg of total tRNA, 1 μL of diluted DMS solution (dilution of saturated stock solution was twofold in 100% ethanol) or 5–10 μL of pure DEPC. Two different conditions were tested: native and semi-denaturing. Under native conditions, DMS modifications were done for 4 min in 50 mM sodium cacodylate, pH 7.5, and 20 mM magnesium acetate for N-3 modification of C residues, or in 50 mM sodium cacodylate, pH 7.2, and 10 mM $MgCl_2$ for N-7 methylations of G residues. Semi-denaturing conditions were as native conditions for incubation times, but the buffer contained 50 mM sodium cacodylate, pH 7.5, and 1 mM EDTA. Under native conditions, DEPC modifications were done in 50 mM sodium cacodylate, pH 7.2, and 10 mM $MgCl_2$, with 10 μL of DEPC for 15 min. Under semi-denaturing conditions, 5 μL of DEPC were applied for 15 min in 50 mM sodium cacodylate, pH 7.5, and 1 mM EDTA. Reactions were stopped and modifications or cleavage sites were assigned as described previously (Felden et al., 1994a).

Computer modeling

Graphical modeling was performed on an Evans and Sutherland PS300 display system. Construction of the model was performed as described for the BMV RNA (Felden et al., 1994a) or more generally as described by Westhof (1993). The graphical model was subjected to restrained least-squares refinement according to Konnert and Hendrickson (1980) with the programs NUCLIN and NUCLSQ (Westhof et al., 1985) in order to ensure correct geometry and stereochemistry. Graphic representations were produced with DRAWNA software (Massire et al., 1994).

ACKNOWLEDGMENTS

We thank J. Witz for supplying us with TMV, A. Hoeft for synthesis of DNA fragments, and B. Masquida for the 2D drawings shown in Figures 1 and 5. This work was supported by the Centre National de la Recherche Scientifique (CNRS), the Ministère de l'Enseignement Supérieur et de la Recherche (MESR), and the Université Louis Pasteur (Strasbourg). B.F. was supported partially by grants from MESR and Association pour la Recherche sur le Cancer.

Received December 13, 1995; returned for revision January 9, 1996; revised manuscript received February 21, 1996

REFERENCES

Avila-Rincon MJ, Ferrero ML, Alonso E, Garcia-Luque I, Diaz-Ruiz JR. 1989. Nucleotide sequence of 5' and 3' non-coding regions of pepper mild mottle virus strain S RNA. *J Gen Virol* 70:3025–3031.

Bloomer AC, Champness JN, Bricogne G, Staden R, Klug A. 1978. Protein disk of tobacco mosaic virus at 2.8 Å showing the interactions within and between subunits. *Nature 276*:362–368.

Brunel C, Romby P, Westhof E, Ehresmann C, Ehresmann B. 1991. Three-dimensional model of *Escherichia coli* ribosomal 5S RNA as deduced from structure probing in solution and computer modelling. *J Mol Biol 221*:293–308.

Duckett DR, Lilley DMJ. 1990. The three-way DNA junction is a Y-shaped molecule in which there is no helix-helix stacking. *EMBO J 9*:1659–1664.

Duckett DR, Murchie AIH, Lilley DMJ. 1995. The global folding of four-way helical junctions in RNA, including that in U1 snRNA. *Cell 83*:1027–1036.

Dumas P, Moras D, Florentz C, Giegé R, Verlaan P, Van Belkum A, Pleij CWA. 1987. 3-D graphics modelling of the tRNA-like 3'-end of turnip yellow mosaic virus RNA: Structural and functional implications. *J Biomol Struct Dyn 4*:707–728.

England TE, Uhlenbeck OC. 1978. Enzymatic oligoribonucleotide synthesis with T4 RNA ligase. *Biochemistry 17*:2069–2076.

Felden B, Florentz C, Giegé R, Westhof E. 1994a. Solution structure of the tRNA-like 3'-end of brome mosaic virus genomic RNAs. Conformational mimicry with canonical tRNAs. *J Mol Biol 235*: 508–531.

Felden B, Florentz C, McPherson A, Giegé R. 1994b. A histidine accepting tRNA-like fold at the 3'-end of satellite tobacco mosaic virus RNA. *Nucleic Acids Res 22*:2882–2886.

Francklyn C, Schimmel P. 1990. Enzymatic aminoacylation of an eight base-pair microhelix with histidine. *Proc Natl Acad Sci USA 87*: 8655–8659.

Frugier M, Florentz C, Giegé R. 1994. Efficient aminoacylation of resected RNA helices by class II aspartyl-tRNA synthetase dependent on a single nucleotide. *EMBO J 13*:2218–2226.

Gallie DR, Feder JN, Schimke RT, Walbot V. 1991. Functional analysis of the tobacco mosaic virus tRNA-like structure in cytoplasmic gene regulation. *Nucleic Acids Res 19*:5031–5036.

Gallie DR, Walbot V. 1992. Identification of the motifs within the tobacco mosaic virus 5'-leader responsible for enhancing translation. *Nucleic Acids Res 20*:4631–4638.

Garcia-Arenal F. 1988. Sequence and structure at the genome 3' end of the U2-strain of tobacco mosaic virus, a histidine-accepting tobamovirus. *Virology 167*:201–206.

Giegé R, Briand JP, Mengual R, Ebel JP, Hirth L. 1978. Valylation of the two components of turnip yellow mosaic virus and specificity of the aminoacylation reaction. *Eur J Biochem 84*:251–256.

Goelet P, Lomonossoff GP, Butler PJG, Akam ME, Gait MJ, Karn J. 1982. Nucleotide sequence of tobacco mosaic virus RNA. *Proc Natl Acad Sci USA 79*:5818–5822.

Guilley H, Jonard G, Kukla B, Richards KE. 1979. Sequence of 1000 nucleotides at the 3' end of tobacco mosaic virus RNA. *Nucleic Acids Res 6*:1287–1308.

Himeno H, Hasegawa T, Ueda T, Watanabe K, Miura K, Shimizu M. 1989. Role of the extra G-C pair at the end of the acceptor stem of $tRNA^{His}$ in aminoacylation. *Nucleic Acids Res 17*:7855–7863.

Holmes KC. 1984. The structure determination of tobacco mosaic virus. In: Jurnak FA, McPherson A, eds. *Biological macromolecules and assemblies, virus structures, vol 1*. New York: John Wiley & Sons. pp 121–148.

Joshi RL, Chapeville F, Haenni AL. 1985. Conformational requirements of tobacco mosaic virus RNA for aminoacylation and adenylation. *Nucleic Acid Res 13*:347–354.

Konnert JH, Hendrickson WA. 1980. A restrained-parameter thermal factor refinement procedure. *Acta Crystallogr A 36*:344–350.

Krol A, Westhof E, Bach M, Luhrmann R, Ebel JP, Carbon P. 1990. Solution structure of human U1 snRNA. Derivation of a possible three-dimensional model. *Nucleic Acids Res 18*:3803–3811.

Leathers V, Tanguay R, Kobayashi M, Gallie DR. 1993. A phylogenetically conserved sequence within viral 3' untranslated RNA pseudoknots regulates translation. *Mol Cell Biol 13*:5331–5347.

Leontis NB, Hills MT, Piotto M, Malhotra A, Nussbaum J, Gorenstein DG. 1993. A model for the solution structure of a branched, three-strand DNA complex. *J Biomol Struct Dyn 11*:215–223.

Lilley DMJ, Clegg RM. 1993. The structure of branched DNA species. *Quart Rev Biophys 26*:131–175.

Mans RMW, Pleij CWA, Bosch L. 1991. tRNA-like structures. Structure, function and biological significance. *Eur J Biochem 201*: 303–324.

Massire C, Gaspin C, Westhof E. 1994. DRAWNA: A program for drawing schematic views of nucleic acids. *J Mol Graphics 12*:201–206.

Mengual R. 1977. Cinétiques d'aminoacylation des RNA des virus de la mosaïque jaune du navet et de la mosaïque du tabac et mise au point de méthodes de purification de la structure aminoacylable [thesis]. Strasbourg, France: Université Louis Pasteur.

Meshi T, Kiyama R, Ohno T, Okada Y. 1983. Nucleotide sequence of the coat protein cistron and the 3' non coding region of cucumber green mottle mosaic virus (water melon strain) RNA. *Virology 127*:54–64.

Namba K, Pattanayek R, Stubbs G. 1989. Visualization of protein-nucleic acid interactions in a virus at 2.9 Å by X-ray fiber diffraction. *J Mol Biol 208*:307–325.

Öberg B, Philipson L. 1972. Binding of histidine to tobacco mosaic virus RNA. *Biochem Biophys Res Commun 48*:927–932.

Ohno T, Aoyagi M, Yamanashi Y, Saito H, Ikawa S, Meshi T, Okada Y. 1984. Nucleotide sequence of the tobacco mosaic virus (tomato strain) genome and comparison with the common strain genome. *J Biochem 96*:1915–1923.

Peattie DA, Gilbert W. 1980. Chemical probes for higher-order structure in RNA. *Proc Natl Acad Sci USA 77*:4679–4682.

Perret V, Garcia A, Puglisi JD, Grosjean H, Ebel JP, Florentz C, Giegé R. 1990. Conformation in solution of yeast $tRNA^{Asp}$ transcripts deprived of modified nucleotides. *Biochimie 72*:735–744.

Perrotta AT, Been MD. 1991. A pseudoknot-like structure required for efficient self-cleavage of hepatitis delta virus RNA. *Nature 19*: 559–564.

Pley HW, Flaherty KM, McKay DB. 1994. Three-dimensional structure of a hammerhead ribozyme. *Nature 372*:6874.

Rich A, RajBhandary UL. 1976. Transfer RNA: Molecular structure, sequence, and properties. *Annu Rev Biochem 45*:805–860.

Rietveld K, Linschooten K, Pleij CWA, Bosch L. 1984. The three-dimensional folding of the tRNA-like structure of tobacco mosaic virus RNA. A new building principle applied twice. *EMBO J 3*:2613–2619.

Rietveld K, Pleij CWA, Bosch L. 1983. Three-dimensional models of the tRNA-like 3'-termini of some plant viral RNAs. *EMBO J 2*: 1079–1085.

Romby P, Moras D, Dumas P, Ebel JP, Giegé R. 1987. Comparison of the tertiary structure of yeast $tRNA^{Asp}$ and $tRNA^{Phe}$ in solution. Chemical modification study of the bases. *J Mol Biol 195*: 193–204.

Rudinger J, Florentz C, Dreher T, Giegé R. 1992. Efficient mischarging of a viral tRNA-like structure and aminoacylation of a minihelix containing a pseudoknot: Histidylation of turnip yellow mosaic virus RNA. *Nucleic Acids Res 20*:1865–1870.

Rudinger J, Florentz C, Giegé R. 1994. Histidylation by yeast HisRS of tRNA ot tRNA-like structure relies on residues −1 and 73 but is dependent on the RNA context. *Nucleic Acids Res 22*:5031–5037.

Scott WG, Finch JT, Klug A. 1995. The crystal structure of an all-RNA hammerhead ribozyme: A proposed mechanism for RNA catalytic cleavage. *Cell 81*:991–1002.

Silberklang M, Prochiantz A, Haenni AL, RajBhandary UL. 1977. Studies on the sequence of the 3' terminal region of turnip yellow mosaic virus. *Eur J Biochem 72*:465–478.

Solis I, Garcia-Arenal F. 1990. The complete nucleotide sequence of the genomic RNA of the tobamovirus tobacco mild green mosaic virus. *Virology 177*:553–558.

Steinberg S, Misch A, Sprinzl M. 1993. Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res 21*:3011–3015.

Stubbs G. 1984. Macromolecular interactions in tobacco mosaic virus. In: Jurnak FA, McPherson A, eds. *Biological macromolecules and assemblies, virus structures, vol 1*. New York: John Wiley & Sons. pp 149–202.

Takamatsu N, Watanabe Y, Meshi T, Okada Y. 1990. Mutational analysis of the pseudoknot region in the 3' non-coding region of tobacco mosaic virus RNA. *J Virol 64*:3686–3693.

Tanner NK, Schaff S, Thill G, Petit-Koskas E, Crain-Denoyelle A-M, Westhof E. 1994. A three-dimensional model of hepatitis delta virus ribozyme based on biochemical and mutational analyses. *Curr Biol 4*:488–498.

ten Dam E, Pleij K, Draper D. 1992. Structural and functional aspects of RNA pseudoknots. *Biochemistry 31*:11665–11676.

Tuschl T, Gohlke C, Jovin TM, Westhof E, Eckstein F. 1994. A three-

dimensional model for the hammerhead ribozyme based on fluorescence measurements. *Science 266*:785–789.
Van Belkum A, Abrahams JP, Pleij CWA, Bosch L. 1985. Five pseudoknots are present at the 204 nucleotides long 3′ non-coding region of tobacco mosaic virus RNA. *Nucleic Acids Res 13*:7673–7686.
Von Wechmar MB, Van Regenmortel MHV. 1970. A simple procedure for purifying tobacco mosaic virus strains. *South African Med J 44*:151–155.
Westhof E. 1993. Modelling the three-dimensional structure of ribonucleic acids. *J Mol Struct (Theochem) 286*:203–210.
Westhof E, Dumas P, Moras D. 1985. Crystallographic refinement of yeast aspartic acid transfer RNA. *J Mol Biol 184*:119–145.
Westhof E, Jaeger L. 1992. RNA pseudoknots: Structural and functional aspects. *Curr Opin Struct Biol 2*:327–333.
Westhof E, Romby P, Romaniuk PJ, Ebel JP, Ehresmann C, Ehresmann B. 1989. Computer modelling from solution data of spinach chloroplast and of *Xenopus laevis* somatic and oocyte 5S rRNAs. *J Mol Biol 207*:417–431.
Wyatt J, Chastain M, Puglisi JD. 1991. Synthesis and purification of large amounts of RNA oligonucleotides. *Biotechniques 11*:764–769.